Markov Chain Review

A sequence of random variables $X_0, X_1, X_2, \cdots$ is called a Markov chain if $\mathbb{P}\{X_n = j \mid X_0 = x_0, X_1 = x_1, \cdots, X_{n-1} = i\}$

$$= \mathbb{P}\{X_n = j \mid X_{n-1} = i\}$$

for all $n \geq 1$ and states i, j.

That is, the chain is governed by one-step transition probabilities, which can be represented by a transition matrix P if the state space is finite.

Powers P^n_{ij} yield the n-step transition prob $\mathbb{P}\{X_n = j \mid X_0 = i\}$

Limiting distribution $\bar{\lambda}$: $\lim_{n\to\infty} P_{ij}^n = \lambda_j$ for all states i & j

(doesn't depend on initial state i)

Equivalently, $\lim_{n\to\infty} P^n = \begin{bmatrix} \bar{\lambda} \\ \vdots \\ \bar{\lambda} \end{bmatrix}$

(all rows converge to $\bar{\lambda}$)

Stationary distribution $\bar{\pi}$: $\bar{\pi} = \bar{\pi} P$

(also called invariant distribution)

$\pi_j = \sum_i \pi_i P_{ij}$ for all states j

Limiting distributions are stationary distributions,

but not vice versa.

Summary of matrix methods for chain computations

① For irreducible, aperiodic MC, the long-term probability of visiting each state is given by $\bar{\pi}$ (left eigenvector for eigenvalue 1)
The first return time for state j is given by $1/\pi_j$

② For a reducible MC, organize transition matrix $P = \begin{array}{c|cc} & t & r \\ \hline t & Q & R \\ r & 0 & \tilde{P} \end{array}$

For transient states, $M = (I-Q)^{-1}$ has entries

M_{ji} = expected # of visits to transient state i if start at transient state j

Sum j^{th} row of M to obtain total expected # of visits to transient states if start at transient state j

③ To apply to irreducible chain, we can compute the expected # of visits to state i before hitting state k if start at j by turning k into an absorbing state (so rest of states become transient). Then use matrix M.

④ For irreducible MC, to find the probability of hitting state b before state c, let b & c be absorbing states.

$$P = \begin{array}{c c} & \begin{array}{cc} & b\ c \end{array} \\ \begin{array}{c} t \\ r \end{array} & \left[\begin{array}{c:c} Q & R \\ \hdashline 0 & I \end{array}\right] \end{array}$$

Let A be the matrix with entries $A_{ij} = \mathbb{P}\{X_n = j \text{ for some } n \geq 1 : X_0 = i\}$

For transient state i and absorbing state j,

$$A_{ij} = \mathbb{P}\{X_n = j \text{ eventually} \mid X_0 = i\}$$

$$= \sum_{\text{all states } k} \mathbb{P}\{X_1 = k \mid X_0 = i\}\, \mathbb{P}\{X_n = j \text{ eventually} \mid X_1 = k\}$$

$$= \sum_{\text{transient } k} \underbrace{\mathbb{P}\{X_1 = k \mid X_0 = i\}}_{Q_{ik}}\, \underbrace{\mathbb{P}\{X_n = j \text{ eventually} \mid X_1 = k\}}_{A_{kj}}$$

$$+ \sum_{\text{recurrent } k} \underbrace{\mathbb{P}\{X_1 = k \mid X_0 = i\}}_{R_{ik}}\, \underbrace{\mathbb{P}\{X_n = j \text{ eventually} \mid X_1 = k\}}_{= \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}}$$

$$= \sum_{\text{transient } k} Q_{ik} A_{kj} + R_{ij} \Rightarrow A = QA + R$$

$$\Rightarrow (I - Q)A = R$$

$$\Rightarrow A = (I - Q)^{-1} R = MR$$

Chapter 4 Branching Processes

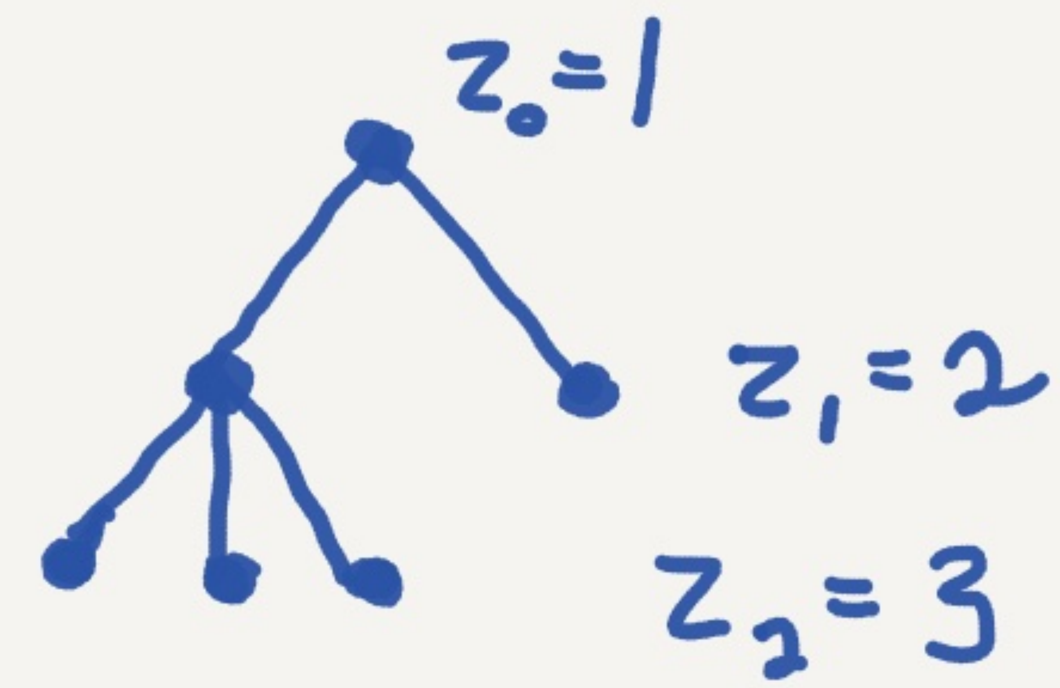
- stochastic model for population growth
- late 1800s, Galton & Watson developed to track male surnames from generation to generation
- historical note: Galton was an influential statistician who created the concept of correlation. He also advocated social Darwinism, eugenics, and scientific racism.

  Watson was a mathematician working on electricity & magnetism.

Def'n: A __branching process__ is a sequence of RVs $Z_0, Z_1, Z_2, \cdots$ with a state space $\{0,1,2,\cdots\}$, where Z_n = # individuals in generation n, in which each individual independently produces offspring according to an offspring distribution $a = (a_0, a_1, a_2, \cdots)$, $0 \le a_k \le 1$, $\sum_{k=0}^{\infty} a_k = 1$,

a_k = prob of k offspring

Typically start with $Z_0 = 1$:



Note that 0 will be an absorbing state (extinction)

What can we say about the population size over time?

If $a_0 = 0$, then always positive. If $a_0 = 1$, instant extinction.
Will usually assume $0 < a_0 < 1$ __and__ $a_0 + a_1 < 1$, so positive prob of multiple offspring.

Lemma 4.1 If $0 < a_0 < 1$ and $a_0 + a_1 < 1$, then all nonzero states are transient.

Proof: Note that prob of going from i to 0 is $(a_0)^i$ all i individuals had 0 offspring

Calculate f_i = prob of eventually hitting i again if start at i

($f_i < 1$ implies i transient)

Section 4.2 Mean generation size

Let X_i = # offspring of i^{th} individual in $(n-1)^{st}$ generation (iid with distribution a)

Note $Z_n = \sum_{i=1}^{Z_{n-1}} X_i$ (sum offspring over the individuals in $(n-1)^{st}$ generation)

Mean of offspring distribution: $\mu = \sum_{k=0}^{\infty} k a_k = \mathbb{E}[X_i]$

Variance is $\sigma^2 = \sum_{k=0}^{\infty} a_k (k-\mu)^2$

Expected value of n^{th} generation is then

$\mathbb{E}[Z_n] =$

So if $Z_0 = 1$, then $\mathbb{E}[Z_n] = \mu^n$.

(In general, $\mathbb{E}[Z_n] = \mu^n \mathbb{E}[Z_0]$.)

Three cases can occur:

1. Subcritical $\mu < 1$

$$\mathbb{P}\{Z_n = 0\} = 1 - \mathbb{P}\{Z_n \geq 1\} \geq 1 - \mathbb{E}[Z_n] = 1 - \mu^n \to 1 \text{ as } n \to \infty$$

2. Critical $\mu = 1$

3. Supercritical $\mu > 1$

$\boxed{\text{R demo}}$

When can extinction occur? When can population explosions occur?

Variance of generation size

$$\mathrm{Var}[Z_n] = \mathrm{Var}[\mathbb{E}[Z_n \mid Z_{n-1}]] + \mathbb{E}[\mathrm{Var}[Z_n \mid Z_{n-1}]]$$ Law of Total Variance

$$= \mathrm{Var}[\underset{\substack{\uparrow \\ \mathbb{E}[a]}}{\mu} Z_{n-1}] + \mathbb{E}[\underset{\substack{\uparrow \\ \mathrm{Var}[a]}}{\sigma^2} Z_{n-1}]$$ ← using independence of X_i

$$= \mu^2 \mathrm{Var}[Z_{n-1}] + \sigma^2 \mu^{n-1}$$

If $Z_0 = 1$, then $\mathrm{Var}[Z_0] = 0$

$$\mathrm{Var}[Z_1] = \sigma^2$$

$$\mathrm{Var}[Z_2] = \sigma^2\mu^2 + \sigma^2\mu = \sigma^2\mu(1+\mu)$$

$$\vdots$$

$$\mathrm{Var}[Z_n] = \sigma^2\mu^{n-1}\sum_{k=0}^{n-1}\mu^k = \begin{cases} n\sigma^2 & \text{if } \mu = 1 \\ \sigma^2\mu^{n-1}\left(\frac{\mu^n - 1}{\mu - 1}\right) & \text{if } \mu \neq 1 \end{cases}$$

(Proof by induction)

So if $Z_0 = 1$, then $\mathbb{E}[Z_n] = \mu^n$ and

$$\mathrm{Var}[Z_n] = \sigma^2 \mu^{n-1} \sum_{k=0}^{n-1} \mu^k = \begin{cases} n\sigma^2 & \text{if } \mu = 1 \\ \sigma^2 \mu^{n-1}\left(\frac{\mu^n - 1}{\mu - 1}\right) & \text{if } \mu \neq 1 \end{cases}$$

① Subcritical $\mu < 1$: $\lim_{n\to\infty} \mathbb{E}[Z_n] = 0$

$\lim_{n\to\infty} \mathbb{P}\{Z_n = 0\} = 1$

$\lim_{n\to\infty} \mathrm{Var}[Z_n] = 0$

② Critical $\mu = 1$: $\lim_{n\to\infty} \mathbb{E}[Z_n] = 1$ but $\mathrm{Var}[Z_n] = n\sigma^2 \to \infty$ as $n \to \infty$

(enormous variability in pop size)

③ Supercritical $\mu > 1$: $\lim_{n\to\infty} \mathbb{E}[Z_n] = \infty$

$\lim_{n\to\infty} \mathrm{Var}[Z_n] = \infty$

(growing exponentially fast as $n \to \infty$)

Can get either extinction or exponential pop growth in cases ② & ③.
What are the probabilities of each, depending on offspring distribution a?